RAW SEQUENCE LISTING DATE: 05/05/2003
PATENT APPLICATION: **US/09/858,332C** TIME: 09:43:03

Input Set : **N:\Crf4\04182003\I858332B.raw**
Output Set: **N:\CRF4\05052003\I858332C.raw**


```
      1 <110> APPLICANT: Tchaga, Grigory S.
      2       Jokhadze, George
      3 <120> TITLE OF INVENTION: Metal Ion Affinity Tags and Methods for
      4       Using the Same
      5 <130> FILE REFERENCE: CLON-056CIP
C-->  6 <140> CURRENT APPLICATION NUMBER: US/09/858,332C
      7 <141> CURRENT FILING DATE: 2003-04-15
      8 <150> PRIOR APPLICATION NUMBER: 09/404,017
      9 <151> PRIOR FILING DATE: 1999-09-23
     10 <150> PRIOR APPLICATION NUMBER: 60/101,867
     11 <151> PRIOR FILING DATE: 1998-09-25
     12 <160> NUMBER OF SEQ ID NOS: 20
     13 <170> SOFTWARE: FastSEQ for Windows Version 4.0
     15 <210> SEQ ID NO: 1
     16 <211> LENGTH: 16
     17 <212> TYPE: PRT
     18 <213> ORGANISM: Artificial Sequence
     19 <220> FEATURE:
     20 <223> OTHER INFORMATION: affinity peptide
     21 <400> SEQUENCE: 1
     22       His Leu Ile His Asn Val His Lys Glu Glu His Ala His Ala His Asn
     23        1               5                  10                  15
     25 <210> SEQ ID NO: 2
     26 <211> LENGTH: 18
     27 <212> TYPE: PRT
     28 <213> ORGANISM: Artificial Sequence
     29 <220> FEATURE:
     30 <223> OTHER INFORMATION: affinity peptide
     31 <400> SEQUENCE: 2
     32       His Asp Asp His Asp Asp His Asp Asp His Asp Asp His Asp Asp His
     33        1               5                  10                  15
     34       Asp Asp
     36 <210> SEQ ID NO: 3
     37 <211> LENGTH: 16
     38 <212> TYPE: PRT
     39 <213> ORGANISM: Artificial Sequence
     40 <220> FEATURE:
     41 <223> OTHER INFORMATION: affinity peptide
     42 <400> SEQUENCE: 3
     43       His Glu Glu His Glu Glu His Glu Glu His Glu Glu His Glu Glu His
     44        1               5                  10                  15
     46 <210> SEQ ID NO: 4
     47 <211> LENGTH: 18
```

RAW SEQUENCE LISTING DATE: 05/05/2003
PATENT APPLICATION: US/09/858,332C TIME: 09:43:03

Input Set : N:\Crf4\04182003\I858332B.raw
Output Set: N:\CRF4\05052003\I858332C.raw

<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: affinity peptide
<400> SEQUENCE: 4
His Asp Glu His Asp Glu His Asp Glu His Asp Glu His Asp Glu His
1 5 10 15
Asp Glu
<210> SEQ ID NO: 5
<211> LENGTH: 18
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: affinity peptide
<400> SEQUENCE: 5
His Glu Asp His Glu Asp His Glu Asp His Glu Asp His Glu Asp His
1 5 10 15
Glu Asp
<210> SEQ ID NO: 6
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: enterokinase cleavage site
<400> SEQUENCE: 6
Ile Glu Gly Arg
1
<210> SEQ ID NO: 7
<211> LENGTH: 6
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: a factor Xa cleavage site
<400> SEQUENCE: 7
Leu Val Pro Arg Gly Ser
1 5
<210> SEQ ID NO: 8
<211> LENGTH: 8
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: a thrombin cleavage site
<400> SEQUENCE: 8
His Pro Phe His Leu Val Ile His
1 5
<210> SEQ ID NO: 9
<211> LENGTH: 10
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence

RAW SEQUENCE LISTING DATE: 05/05/2003
PATENT APPLICATION: **US/09/858,332C** TIME: 09:43:03

Input Set : **N:\Crf4\04182003\I858332B.raw**
Output Set: **N:\CRF4\05052003\I858332C.raw**

<220> FEATURE:
<223> OTHER INFORMATION: a renin cleavage site
<400> SEQUENCE: 9
Cys Tyr Pro Tyr Asp Val Pro Asp Tyr Ala
1 5 10
<210> SEQ ID NO: 10
<211> LENGTH: 8
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: an immunological tag
<400> SEQUENCE: 10
Asp Tyr Lys Asp Asp Asp Asp Lys
1 5
<210> SEQ ID NO: 11
<211> LENGTH: 11
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: an immunological tag
<400> SEQUENCE: 11
Cys Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu
1 5 10
<210> SEQ ID NO: 12
<211> LENGTH: 11
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: an immunological tag
<400> SEQUENCE: 12
Cys Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu
1 5 10
<210> SEQ ID NO: 13
<211> LENGTH: 3430
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: DNA sequence of vector containing cDNA of recombinant enterokinase
<400> SEQUENCE: 13
gacgaaaggg cctcgtgata cgcctatttt ataggttaa tgtcatgata ataatggttt 60
cttagacgtc aggtggcact tttcggggaa atgtgcgcgg aacccctatt tgtttatttt 120
tctaaataca ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat 180
aatattgaaa aaggaagagt atgagtattc aacatttccg tgtcgccctt attccctttt 240
ttgcggcatt ttgccttcct gtttttgctc acccagaaac gctggtgaaa gtaaaagatg 300
ctgaagatca gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga 360
tccttgagag ttttcgcccc gaagaacgtt ttccaatgat gagcactttt aaagttctgc 420
tatgtggcgc ggtattatcc cgtattgacg ccgggcaaga gcaactcggt cgccgcatac 480
actattctca gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg 540

RAW SEQUENCE LISTING DATE: 05/05/2003
PATENT APPLICATION: US/09/858,332C TIME: 09:43:03

Input Set : N:\Crf4\04182003\I858332B.raw
Output Set: N:\CRF4\05052003\I858332C.raw

155 gcatgacagt aagagaatta tgcagtgctg ccataaccat gagtgataac actgcggcca 600
156 acttacttct gacaacgatc ggaggaccga aggagctaac cgctttttg cacaacatgg 660
157 gggatcatgt aactcgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaacg 720
158 acgagcgtga caccacgatg cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg 780
159 gcgaactact tactctagct tcccggcaac aattaataga ctggatggag gcggataaag 840
160 ttgcaggacc acttctgcgc tcggcccttc cggctggctg gtttattgct gataaatctg 900
161 gagccggtga gcgtgggtct cgcggtatca ttgcagcact ggggccagat ggtaagccct 960
162 cccgtatcgt agttatctac acgacgggga gtcaggcaac tatggatgaa cgaaatagac 1020
163 agatcgctga gataggtgcc tcactgatta agcattggta actgtcagac caagtttact 1080
164 catatatact ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga 1140
165 tccttttga taatctcatg accaaaatcc cttaacgtga gttttcgttc cactgagcgt 1200
166 cagaccccgt agaaaagatc aaaggatctt cttgagatcc ttttttttctg cgcgtaatct 1260
167 gctgcttgca aacaaaaaaa ccaccgctac cagcggtggt ttgtttgccg gatcaagagc 1320
168 taccaactct ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgtcc 1380
169 ttctagtgta gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc 1440
170 tcgctctgct aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg 1500
171 ggttggactc aagacgatag ttaccggata aggcgcagcg gtcgggctga acgggggggtt 1560
172 cgtgcacaca gcccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg 1620
173 agctatgaga aagcgccacg cttcccgaag ggagaaaggc ggacaggtat ccggtaagcg 1680
174 gbgcagggtc ggaacaggag agcgcacgag ggagcttcca gggggaaacg cctggtatct 1740
175 ttatagtcct gtcgggtttc gccacctctg acttgagcgt cgatttttgt gatgctcgtc 1800
176 agggggggcgg agcctatgga aaaacgccag caacgcggcc tttttacggt tcctggcctt 1860
177 ttgctggcct tttgctcaca tgttctttcc tgcgttatcc cctgattctg tggataaccg 1920
178 tattaccgcc tttgagtgag ctgataccgc tcgccgcagc cgaacgaccg agcgcagcga 1980
179 gtcagtgagc gaggaagcgg aagagcgccc aatacgcaaa ccgcctctcc ccgcgcgttg 2040
180 gccgattcat taatgcagct ggcacgacag gtttcccgac tggaaagcgg gcagtgagcg 2100
181 caacgcaatt aatgtgagtt agctcactca ttaggcaccc caggctttac actttatgct 2160
182 tccggctcgt atgttgtgtg gaattgtgag cggataacaa tttcacacag gaaacagcta 2220
183 tgaccatgat tacgccaagc ttgaaggatc atctcatcca caatgtccac aaagaggagc 2280
184 acgctcatgc ccacaacaag atcgatattg tcggaggaag tgactccaga gaaggagcct 2340
185 ggccttgggt cgttgctctg tatttcgacg atcaacaggt ctgcggagct tctctggtga 2400
186 gcagggattg gctggtgtcg gccgcccact gcgtgtacgg gagaaatatg gagccgtcta 2460
187 agtggaaagc agtgctaggc ctgcatatgg catcaaatct gacttctcct cagatagaaa 2520
188 ctaggttgat tgaccaaatt gtcataaacc cacactacaa taaacggaga aagaacaatg 2580
189 acattgccat gatgcatctt gaaatgaaag tgaactacac agattatata cagcctattt 2640
190 gtttaccaga agaaaatcaa gtttttcccc caggaagaat ttgttctatt gctggctggg 2700
191 gggcacttat atatcaaggt tctactgcag acgtactgca agaagctgac gttccccttc 2760
192 tatcaaatga gaaatgtcaa caacagatgc cagaatataa cattacggaa aatatggtgt 2820
193 gtgcaggcta tgaagcagga ggggtagatt cttgtcaggg ggattcaggc ggaccactca 2880
194 tgtgccaaga aaacaacaga tggctcctgg ctggcgtgac gtcatttgga tatcaatgtg 2940
195 cactgcctaa tcgcccaggg gtgtatgccc gggtcccaag gttcacagag tggatacaaa 3000
196 gttttctaca tgagctcgta attagctgag aattcactgg ccgtcgtttt acaacgtcgt 3060
197 gagcctggga aaaccctggc gttacccaac ttaatcgcct tgcagcacat ccccctttcg 3120
198 ccagctggcg taatagcgaa gaggcccgca ccgatcgccc ttcccaacag ttgcgcagcc 3180
199 tgaatggcga atggcgcctg atgcggtatt ttctccttac gcatctgtgc ggtatttcac 3240
200 accgcatatg gtgcactctc agtacaatct gctctgatgc cgcatagtta agccagcccc 3300
201 gacacccgcc aacacccgct gacgcgccct gacgggcttg tctgctcccg gcatccgctt 3360
202 acagacaagc tgtgaccgtc tccgggagct gcatgtgtca gaggttttca ccgtcatcac 3420
203 cgaaacgcgc 3430

RAW SEQUENCE LISTING DATE: 05/05/2003
PATENT APPLICATION: US/09/858,332C TIME: 09:43:03

Input Set : N:\Crf4\04182003\I858332B.raw
Output Set: N:\CRF4\05052003\I858332C.raw

<210> SEQ ID NO: 14
<211> LENGTH: 279
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: protein sequence of vector containing cDNA of recombinant enterokinase
<400> SEQUENCE: 14

```
Met Thr Met Ile Thr Pro Ser Leu Lys Asp His Leu Ile His Asn Val
 1               5                  10                  15
His Lys Glu Glu His Ala His Ala His Asn Lys Ile Asp Ile Val Gly
            20                  25                  30
Gly Ser Asp Ser Arg Glu Gly Ala Trp Pro Trp Val Val Ala Leu Tyr
        35                  40                  45
Phe Asp Asp Gln Gln Val Cys Gly Ala Ser Leu Val Ser Arg Asp Trp
    50                  55                  60
Leu Val Ser Ala Ala His Cys Val Tyr Gly Arg Asn Met Glu Pro Ser
65                  70                  75                  80
Lys Trp Lys Ala Val Leu Gly Leu His Met Ala Ser Asn Leu Thr Ser
                85                  90                  95
Pro Gln Ile Glu Thr Arg Leu Ile Asp Gln Ile Val Ile Asn Pro His
            100                 105                 110
Tyr Asn Lys Arg Arg Lys Asn Asn Asp Ile Ala Met Met His Leu Glu
        115                 120                 125
Met Lys Val Asn Tyr Thr Asp Tyr Ile Gln Pro Ile Cys Leu Pro Glu
    130                 135                 140
Glu Asn Gln Val Phe Pro Pro Gly Arg Ile Cys Ser Ile Ala Gly Trp
145                 150                 155                 160
Gly Ala Leu Ile Tyr Gln Gly Ser Thr Ala Asp Val Leu Gln Glu Ala
                165                 170                 175
Asp Val Pro Leu Leu Ser Asn Glu Lys Cys Gln Gln Gln Met Pro Glu
            180                 185                 190
Tyr Asn Ile Thr Glu Asn Met Val Cys Ala Gly Tyr Glu Ala Gly Gly
        195                 200                 205
Val Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Met Cys Gln Glu
    210                 215                 220
Asn Asn Arg Trp Leu Leu Ala Gly Val Thr Ser Phe Gly Tyr Gln Cys
225                 230                 235                 240
Ala Leu Pro Asn Arg Pro Gly Val Tyr Ala Arg Val Pro Arg Phe Thr
                245                 250                 255
Glu Trp Ile Gln Ser Phe Leu His Glu Leu Val Ile Ser Glu Phe Thr
            260                 265                 270
Gly Arg Arg Phe Thr Thr Ser
        275
```

<210> SEQ ID NO: 15
<211> LENGTH: 12
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:

VERIFICATION SUMMARY DATE: 05/05/2003
PATENT APPLICATION: **US/09/858,332C** TIME: 09:43:04

Input Set : **N:\Crf4\04182003\I858332B.raw**
Output Set: **N:\CRF4\05052003\I858332C.raw**

L:6 M:270 C: Current Application Number differs, Wrong Format